ZooKeys 911: 139–160 (2020)
doi: 10.3897/zookeys.911.39222
http://zookeys.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
ZooKeys
Launched to accelerate biodiversity research

Genetic diversity and population structure of *Terapon jarbua* (Forskål, 1775) (Teleostei, Terapontidae) in Malaysian waters

Shyama Sundari Devi Chanthran[1,2], Phaik-Eem Lim[1], Yuan Li[3], Te-Yu Liao[4], Sze-Wan Poong[1], Jianguo Du[3,5,6], Muhammad Ali Syed Hussein[7], Ahemad Sade[8], Richard Rumpet[9], Kar-Hoe Loh[1]

1 *Institute of Ocean and Earth Sciences, University of Malaya, Kuala Lumpur 50603, Malaysia* **2** *Institute for Advanced Studies, University of Malaya, Kuala Lumpur 50603, Malaysia* **3** *Third Institute of Oceanography, Ministry of Natural Resources, Xiamen 361005, China* **4** *Department of Oceanography, National Sun Yat-sen University, Kaohsiung 80424, Taiwan* **5** *Fujian Provincial Station for Field Observation and Research of Island and Coastal Zone in Zhangzhou, Xiamen 361005, China* **6** *Fujian Provincial Key Laboratory of Marine Ecological Conservation and Restoration, Xiamen 361005, China* **7** *Endangered Marine Species Research Unit, Borneo Marine Research Institute, Universiti Malaysia Sabah, Kota Kinabalu 88400, Sabah, Malaysia* **8** *Department of Fisheries Sabah, Kota Kinabalu 88624, Sabah, Malaysia* **9** *Fisheries Research Institute, Sarawak, Department of Fisheries Malaysia, Kuching 93744, Sarawak, Malaysia*

Corresponding author: *Kar-Hoe Loh* (khloh@um.edu.my)

Academic editor: *Nina Bogutskaya* | Received 19 October 2019 | Accepted 2 January 2020 | Published 12 February 2020

http://zoobank.org/66CEE714-50BE-4C33-AF5D-B02A2E2C7DF7

Citation: Chanthran SSD, Lim P-E, Li Y, Liao T-Y, Poong S-W, Du J, Hussein MAS, Sade A, Rumpet R, Loh K-H (2020) Genetic diversity and population structure of *Terapon jarbua* (Forskål, 1775) (Teleostei, Terapontidae) in Malaysian waters. ZooKeys 911: 139–160. https://doi.org/10.3897/zookeys.911.39222

Abstract

A background study is important for the conservation and stock management of a species. *Terapon jarbua* is a coastal Indo-Pacific species, sourced for human consumption. This study examined 134 samples from the central west and east coasts of Peninsular (West) Malaysia and East Malaysia. A 1446-bp concatenated dataset of mtDNA COI and Cyt *b* sequences was used in this study and 83 haplotypes were identified, of which 79 are unique haplotypes and four are shared haplotypes. Populations of *T. jarbua* in Malaysia are genetically heterogenous as shown by the high level of haplotype diversity ranging from 0.9167–0.9952, low nucleotide diversity ranging from 0.0288–0.3434, and high F_{ST} values (within population genetic variation). Population genetic structuring is not distinct as shown by the shared haplotypes between geographic populations and mixtures of haplotypes from different populations within the same genetic cluster. The gene flow pat-

Copyright *Shyama Sundari Devi Chanthran et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

terns and population structuring observed among these regions are likely attributed to geographical distance, past historical events, allopatric speciation, dispersal ability and water currents. For instance, the mixture of haplotypes revealed an extraordinary migration ability of *T. jarbua* (>1200 km) via ancient river connectivity. The negative overall value of the neutrality test and a non-significant mismatch distribution are consistent with demographic expansion(s) in the past. The median-joining network concurred with the maximum likelihood haplotype tree with three major clades resolved. The scarcity of information on this species is an obstacle for future management and conservation purposes. Hence, this study aims to contribute information on the population structure, genetic diversity, and historical demography of *T. jarbua* in Malaysia.

Keywords
COI, crescent perch, Cyt *b*, historical demography, ikan mengkerong, Pleistocene

Introduction

A population's genetic structure describes the total genetic diversity in the population, which is shaped by several factors, including the life history, geographical barriers, gene flow, selection and bottlenecks (Wright 1931; Slatkin 1987; Charlesworth 2009; Liu et al. 2019). The patterns of genetic diversity and population structure provide information on the life histories, demography, reproduction and ecology of a species. This information is important for a population's sustainability by implementation of appropriate conservation and management strategies (Okumuş and Çiftci 2003; Zaya et al. 2017).

Terapon jarbua (Forskål, 1775) is a medium-sized fish commonly known as crescent perch, and it is locally known as "ikan mengkerong" in Malaysia (Department of Fisheries Malaysia 2009). This species is classified under the class Actinopterygii, order Perciformes and family Terapontidae (Froese and Pauly 2018). Although it is primarily a marine species, it has also been found in coastal areas, estuaries, freshwaters and in some coastal lagoons (Rao et al. 2000). It is categorized as a catadromous fish, in which the adults spawn in deeper saltwater while the juveniles move to the shallow sandy bottom area near the river mouths. According to Lavergne et al. (2012), the pelagic larval phase of this species is about 25 days. *Terapon jarbua* is classified as least concern (LC) under the IUCN Red List due to its widespread distribution with no known threats (Dahanukar et al. 2017). The native distributional ranges of the crescent grunters include Australia, Bangladesh, Cambodia, China, India, Indonesia, Japan, Malaysia (Du et al. 2019; Shyama et al. 2020), Mediterranean (Golani and Appelbaum-Golani 2010), Myanmar, Philippines, Red Sea, Sri Lanka and Taiwan (Froese and Pauly 2018).

Existing reports on *T. jarbua* are generally limited to their morphometry (e.g., length-weight relationship) and reproductive biology (Miu et al. 1990; Nandikeswari et al. 2014; Musarrat and Masood 2015) while the information on their genetic diversity and population structures is relatively little. Lavergne et al. (2012) and Liu et al. (2015) reported on the population genetic diversity of *T. jarbua* in the Gulf of Aden, Yemen and Taiwanese waters, respectively, using cytochrome *c* oxidase subunit I (COI), cytochrome *b* (Cyt *b*) and microsatellites molecular markers. Also, a phylogeo-

graphic survey of *T. jarbua* along with other reef fauna of the western Indian Ocean was reported by Borsa et al. (2016) using COI gene sequences. Mitochondrial DNA (mtDNA) has been widely utilized as the marker of choice to examine the genetic diversity and population structure of marine fishes due to its strict maternal inheritance, rapid mutation rates and the absence of recombination in most species (Whitehead et al. 2003; Dowling et al. 2008; Song et al. 2013).

The main focus of the current study is on the Malaysian populations: Peninsular (West) Malaysia and East Malaysia (Sabah and Sarawak) which are located in the tropical Indo-west Pacific region (Fig. 1). These two land masses are about 1200 km apart, separated by the south-western portion of the South China Sea. In this study COI and Cyt *b* were used as molecular markers to examine the level of gene flow, population genetic differentiation and the historical demography of *T. jarbua* populations in Malaysia. To the best of our knowledge, there is no documented report on the population genetics of *T. jarbua* in Malaysia to date. Hence, this study aims to provide a documented background report as well as to fill the information gap for *T. jarbua* in this region. Homologous COI+ Cyt *b* sequences of four regional representatives of this species from India, Taiwan, Hainan and the Philippines were included in the analysis to provide a wider coverage of the species' natural distribution.

Materials and methods

Sampling

Sampling around major landing sites and local markets was conducted in both East and Peninsular Malaysia where 134 samples of various sizes were collected randomly from five wild populations of *T. jarbua*. Populations were provisionally divided into five groups according to region: 1) Kuala Selangor (KS, $N = 31$) of west Peninsular which is surrounded by the Straits of Malacca; 2) Kuantan, Pahang (KN, $N = 30$) of east Peninsular which is adjacent to the South China Sea; 3) Mukah, Sarawak (MH, $N = 21$) of East Malaysia which is surrounded by the South China Sea; 4) Sandakan (SN, $N = 28$) and 5) Tawau (TW, $N = 24$) of East Malaysia which are surrounded by the Sulu Sea and the Celebes Sea, respectively (Fig. 1). Samples were collected within the period of April 2015 to August 2018. Approximately 20 mg of muscle tissue from each fish sample was removed and immediately preserved in 95% ethanol and stored at -20°C until genetic analysis was performed.

DNA extraction, PCR amplification and DNA sequencing

Genomic DNA was extracted using 10% Chelex Resin following the protocol of Hyde et al. (2005). Approximately 680 bp of the COI-5' gene was amplified using the FishF1 or FishF2 forward primers and FishR1 or FishR2 reverse primer pairs (Ward et al. 2005).

Polymerase Chain Reaction (PCR) amplification of approximately 1000 bp from the 5'-end of the Cyt *b* gene was performed using the primer pairs Glu31 and Thr33 (Liu et al. 2015) and internal primers: Glu231 (5'-CTT ACA GGC CTC TTT CTG GCC AT- 3') and Thr233 (5'- TTT GAG CTA CTA ATG CAG TAT- 3') were designed for this study.

PCR was performed using a Mastercycler epgradient S thermalcycler (Eppendorf, Hamburg, Germany) and 25µl reaction mixtures consisting of 12.5µl exTEN 2X PCR master mix (1st BASE, Selangor, Malaysia), 9.5 µl of sterile distilled water, 1µl each of forward and reverse primers, and 1µl of DNA template. PCR cycling conditions were as follow: initial denaturation for 1 min at 96°C, 36 cycles of denaturation at 95°C for 30 s, annealing for 30 s at 44°C (CO1) or 48°C (Cyt *b*), elongation for 1 min at 72°C, and final elongation for 10 min at 72°C. The amplicons were checked for correct length via electrophoresis on a 1% agarose gel (90V for 25 min). PCR products were sent to Apical Scientific Sdn. Bhd. (Selangor, Malaysia) for purification and DNA sequencing.

Sequence analysis

Multiple sequence alignment was first performed separately for each gene region using the CLUSTAL X (Thompson et al. 1994) program implemented in BIOEDIT ver. 7.0.5 (Hall et al. 2011). The sequences were subsequently trimmed and aligned manually prior to concatenation of COI and Cyt *b* sequences. Analyses performed in this study were based on the final truncated length of 1446-bp concatenated sequences. All haplotype sequences were deposited in Genbank under the accession numbers MN529663–MN52993.

Unique haplotypes were quantified and the genetic diversity, nucleotide diversity, and pairwise distance were calculated using DNASP v. 4.0 (Rozas et al. 2003). The level of gene flow among populations (Nm) based on Hudson et al. (1992) was also calculated in DNASP v. 4.0. Analysis of molecular variance (AMOVA) was performed using ARLEQUIN v.3.5 (Excoffier and Lischer 2010) for the four, hypothetical, region-based groupings (Selangor, Pahang, Sabah and Sarawak) to investigate the partition of genetic variation among regions (F_{CT}), among populations within regions (F_{SC}), and within populations (F_{ST}). The significance of the F-statistics for population comparisons was assessed using 1000 permutations. The Tamura Nei plus gamma rate model (TN93+G) was selected by MEGA v. 7.0 (Kumar et al. 2016) as the best-fitting substitution model based on the Bayesian information criterion. A Maximum Likelihood (ML) tree was reconstructed in MEGA 7.0 to show the level of divergence and relationships among haplotypes of *T. jarbua*. The confidence level at each node was assessed by 1000 bootstrap replications. This tree was compared against the median-joining network generated using the program's default settings of NETWORK 4.5.0.2 (Bandelt et al. 1999).

In addition, a neutrality test of the pairwise differences among all populations was performed to infer historical demographic and deviation of sequence variation from evolutionary neutrality. Deviations from neutrality were evaluated using Fu's Fs (Fu 1997) and Tajima's *D* (Tajima 1989) via DNASP. Statistical tests and

confidence intervals for *D* and F's were based on a coalescent simulation algorithm. A large negative value of Fu's Fs or the Tajima's *D* rejecting the null hypothesis of neutrality indicates population expansion(s). The demographic changes were also examined using the mismatch distribution analysis (Rogers and Harpending 1992) in ARLEQUIN with 1000 permutations. The Harpending's raggedness index (Harpending et al. 1993) and the sum of squared deviations (SSD) between observed and expected mismatch for each of the populations under the model of constant population size were analyzed according to Schneider and Excoffier (1999). This method quantifies the smoothness of the observed mismatch distribution and a non-significant result indicates an expanding population. The spatial expansion hypothesis (both raggedness index and SSD) was tested using a parametric bootstrap approach with 1000 replicates.

Results

Genetic diversity

The 1446 bp concatenated COI (631 bp) and Cyt *b* (815 bp) sequences were analyzed for 134 individuals obtained in five different locations (Fig. 1) from East Malaysia and Peninsular Malaysia. The nucleotide composition was 23.0% adenine,

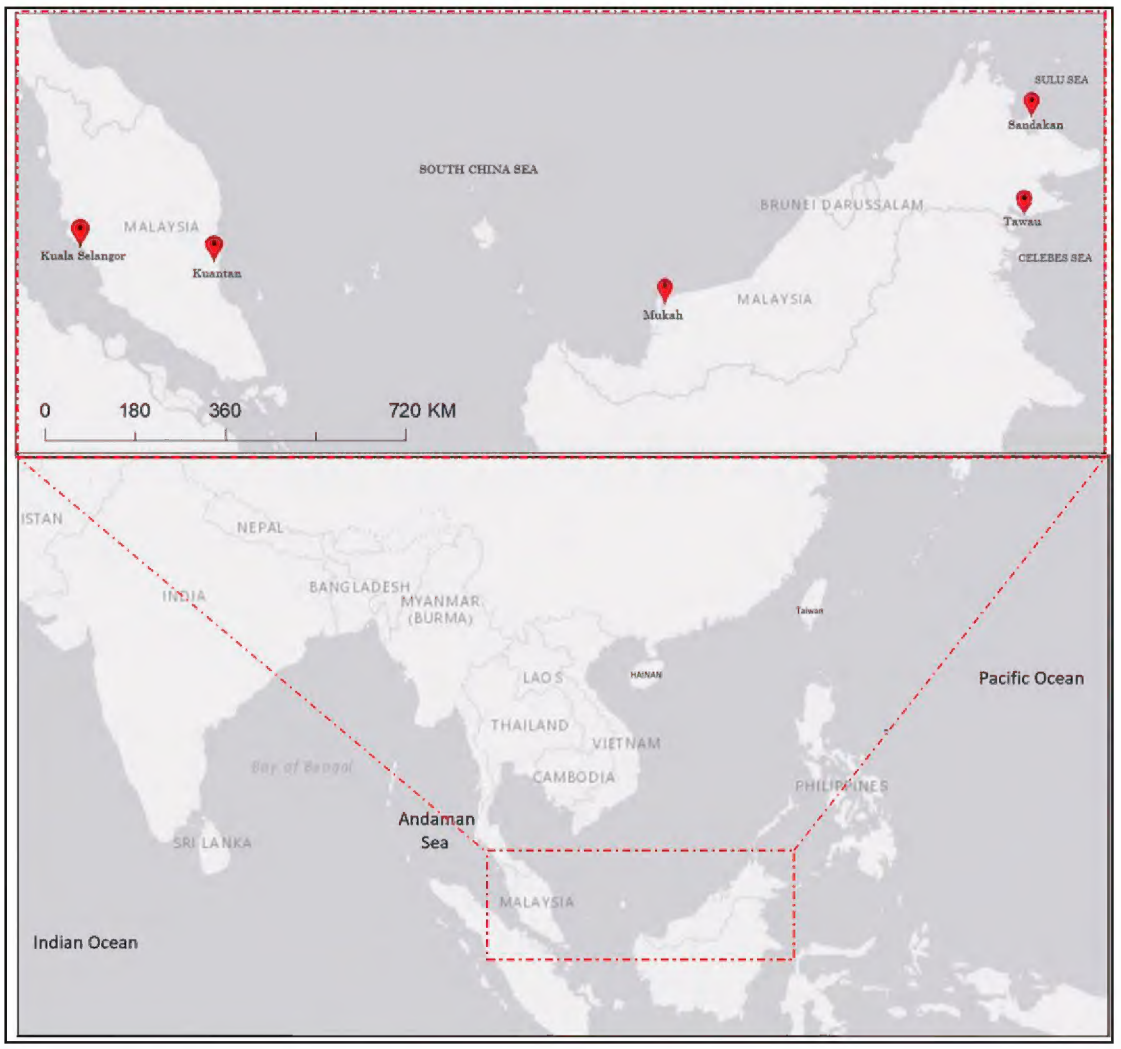

Figure 1. Sampling localities from East (Sandakan and Tawau, Sabah & Mukah, Sarawak) and West (Peninsula) Malaysia (Kuala Selangor, Selangor and Kuantan, Pahang).

28.9% thymine, 32.0% cytosine and 16.1% guanine. The higher A+T content (51.9%) compared to G+C content (48.1%), is common among fishes (Appendix 1). There were 202 polymorphic sites of which 31 (15.4%) were singleton variable sites, and 171 (84.6%) were parsimony informative (Appendix 2). A total of 149 mutations with 66 transitions, 8 transversions and 75 substitutions were found within the dataset (data not shown).

A total of 83 putative haplotypes were derived from the 134 individuals sequenced with 79 of them being unique haplotypes (95.18%) and four were shared haplotypes (4.82%). The dominant haplotype of Malaysian populations is Hap5 (KS, KN, TW, SN, MH, TAI) while other shared haplotypes are Hap3 (KS, KN, TW, MH, IND), Hap8 (KS and KN) and Hap45 (TW and SN). The population from KS recorded the highest total number of haplotypes (22) of which 19 were unique haplotypes, while Tawau recorded the lowest number of haplotypes (15) with 12 unique haplotypes. The nucleotide diversity (π) of *T. jarbua* populations in this study ranged from 0.0288 ±0.0158 (mean ±SD) to 0.3434 ±0.1722 while haplotype diversity (h) ranged from 0.9167 ±0.0482 to 0.9952 ±0.0165 (Table 1). The MH population recorded the highest π and h. A high h and low π indicate that the populations studied were moderate in genetic diversity.

Genetic structure

A ML tree was reconstructed based on the 83 haplotypes of this study and four COI + Cyt *b* sequences from Hainan, Taiwan, India and Philippines which were downloaded from National Center for Biotechnology Information (NCBI) (Appendix 3). The mtDNA concatenated dataset defined the haplotypes into three major clades with no significant clusters corresponding to sampling localities (Fig. 2). Apart from Clade I which consists of eight haplotypes solely from Sarawak, the other four clades include mixtures of haplotypes from various localities without any obvious geographical structuring among them. Clade II consists of haplotypes from MH and KS populations while Clade III consists of the haplotype from the Philippines and haplotypes from Malaysia except MH. Clade III is the most geographically inclusive with haplotypes from India, Hainan, Taiwan, Philippines and all five Malaysian populations.

The general topology of the median-joining network (Fig. 3) corresponded with the ML tree (Fig. 2) with three major clusters identified. In the network presented, shared haplotypes occupy the central area, while the unique haplotypes branched out from the center. This formation provides a star-like profile, which indicates population expansions. Hap3 and Hap5 are the dominant haplotypes in cluster III. The distribution frequency of all the 83 haplotypes in *T. jarbua* populations is presented in Appendix 4. Hap5 recorded the highest distribution frequency with 21 individuals. It is the only common haplotype shared between all five populations from Malaysia and the representative from Taiwan. Hap3 with 16 individuals was found in India and in all locations sampled in Malaysia except for Sandakan.

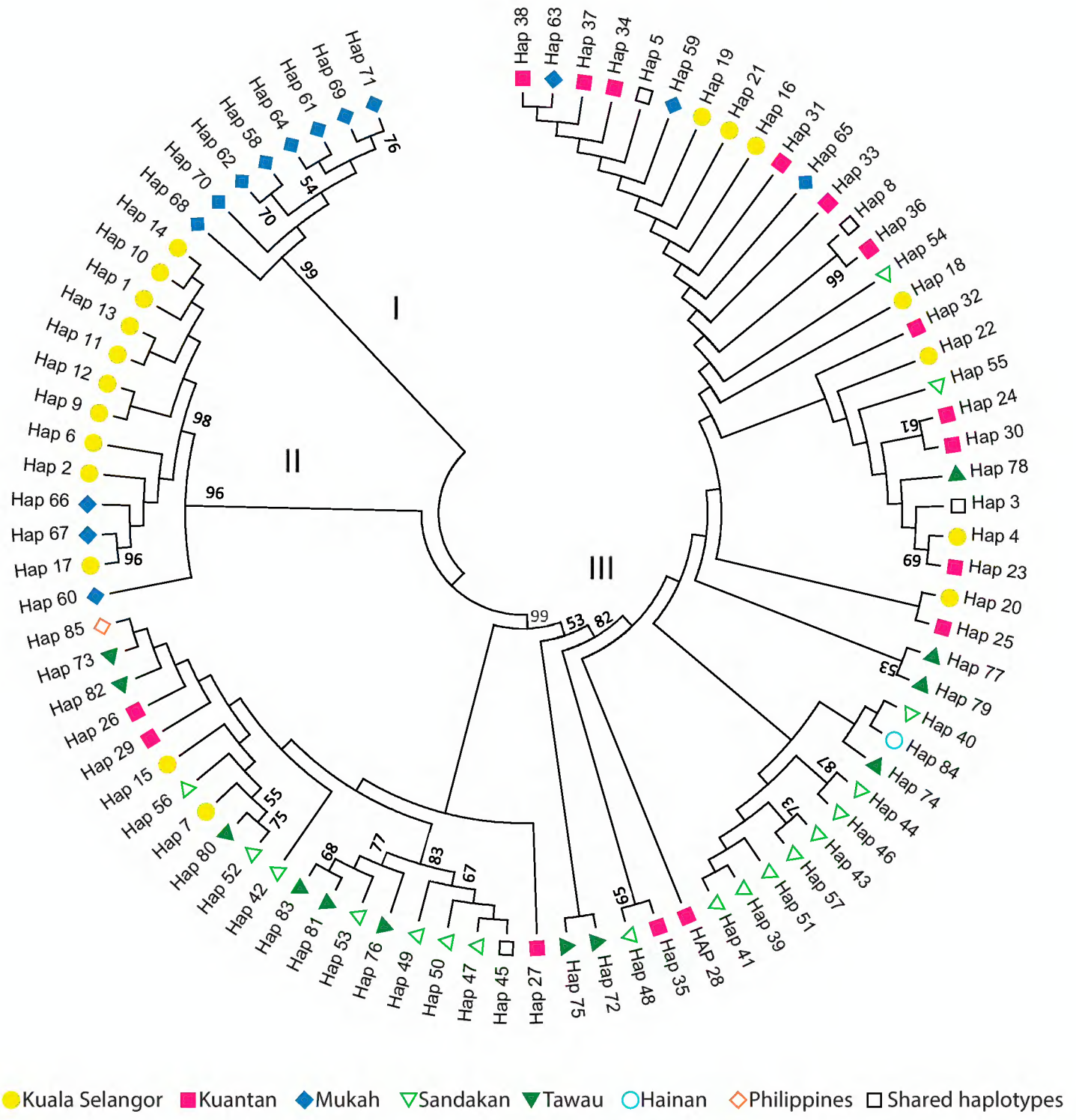

Figure 2. Maximum likelihood haplotype tree reconstructed based on the concatenated mtDNA dataset. The bootstrap values higher than 50% are shown near the nodes.

Table 1. Information and molecular indices of *T. jarbua*. N, number of samples; NH, number of haplotypes; NUH, number of unique haplotypes; *h*, haplotype diversity; π, nucleotide diversity; *k*, average number of pairwise differences.

ID	Populations	N	NH	NUH	*h*	π	*k*
KS	Kuala Selangor, Selangor	31	22	19	0.9828 ±0.0135	0.1817 ±0.0904	36.5161 ±16.3285
KN	Kuantan, Pahang	30	19	16	0.9678 ±0.0208	0.0288 ±0.0158	5.7885 ±2.8485
MH	Mukah, Sarawak	21	16	14	0.9952 ±0.0165	0.3434 ±0.1722	69.0238 ±31.0005
SN	Sandakan, Sabah	28	20	18	0.9577 ±0.0262	0.0487 ±0.0256	9.7810 ±4.6212
TW	Tawau, Sabah	24	15	12	0.9167 ±0.0482	0.0514 ±0.0271	10.3333 ±4.8857
Total		134	83	79	0.9820 ±0.0050	0.0248 ±0.0031	35.8653 ±12.2638

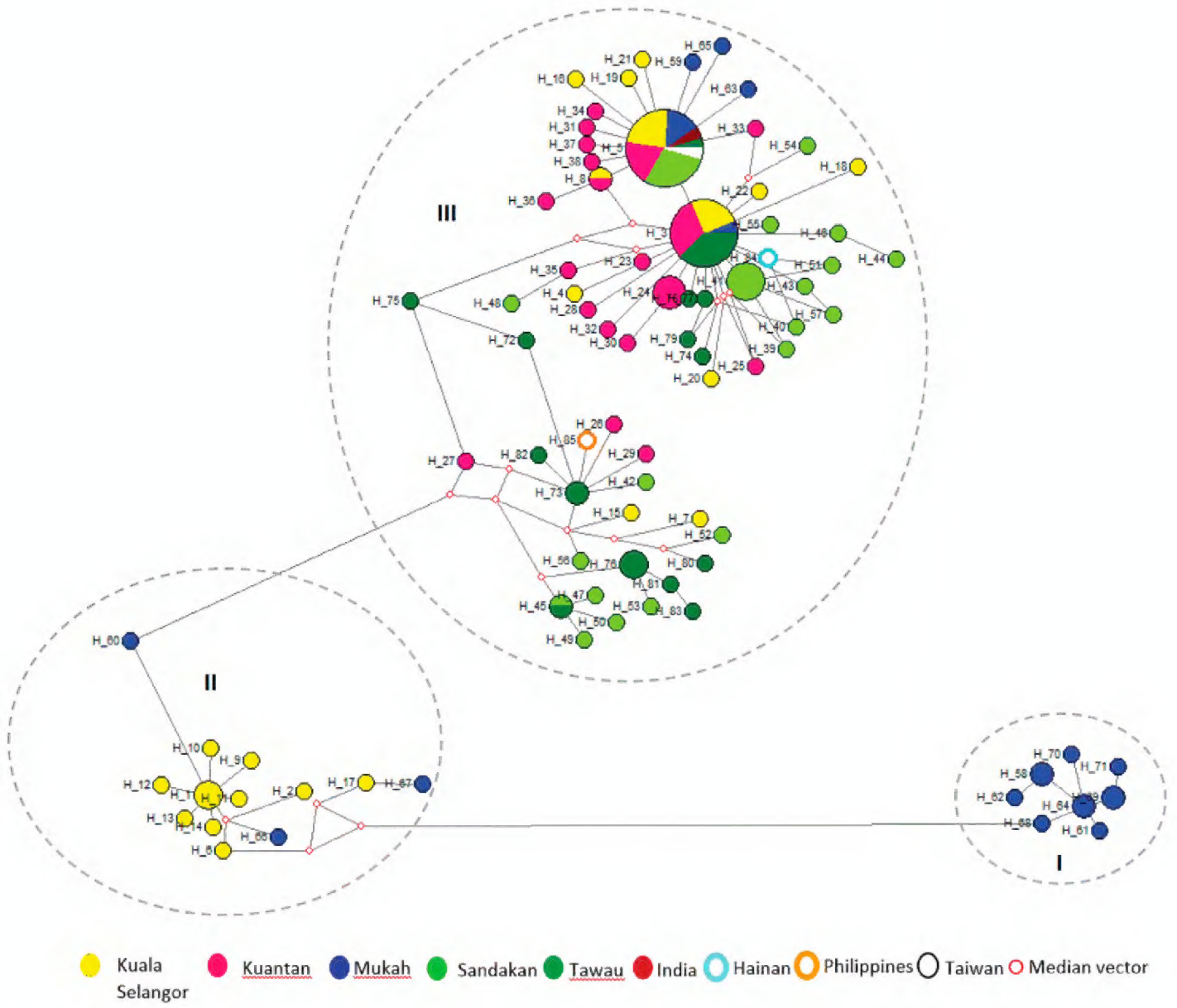

Figure 3. Haplotypes median-joining network corresponding to the ML tree with three observed clusters. The star-like profile observed in cluster III indicates the presence of sudden expansion.

Pairwise F_{ST} comparisons between populations in Malaysia were significant at the 95% confidence level except for the comparison between TW and SN (Table 2). Populations of MH and KN showed the greatest pairwise differentiation (F_{ST} = 0.5353; $p < 0.05$) while SN and TW showed the least differentiation (F_{ST} = 0.0452; $p > 0.05$). The pairwise nucleotide divergence among populations (Table 3) showed the same trend as the F_{ST} values and was not correlated with geographical distance. The overall gene flow (Nm) estimated among populations was low at 0.82. The sequence divergence was calculated using the Kimura two parameter (K2P) distance model for both genes (Table 3). The greatest genetic differences (COI: 0.019 and Cyt *b*: 0.029) were observed between MH-KN, MH-SN and MH-TW. The *T. jarbua* populations displayed a low level of conspecific divergence within 2% (COI).

The genetic structure of the *T. jarbua* populations analysed by AMOVA showed little (39.52%) genetic differentiation among regions but high (62.13%) variation within populations (Table 4). This indicates that the populations were not genetically differentiated among regions and the genetic variation was mainly from within the population level. There is essentially no genetic structuring (-0.14% variation) among populations within region.

Table 2. Pairwise F_{ST} (below diagonal) and exact *P*-values (above diagonal) among five populations of *T. jarbua* based on 1000 permutations of the sequence data set. Numbers in bold represent the highest and lowest value. *Significant at p <0.05 by the permutation test. Overall gene flow (N_m) is 0.82.

Populations	KS	KN	MH	SN	TW
KS	-	0.0000*	0.0000*	0.0000*	0.0000*
KN	0.2965	-	0.0000*	0.0270*	0.0090*
MH	0.3310	**0.5353**	-	0.0000*	0.0000*
SN	0.2681	0.0702	0.5038	-	0.0541
TW	0.2633	0.1773	0.4844	**0.0452**	-

Table 3. Net between-group mean distances using Kimura-2-parameter (K2P) model.

	Populations	KS	KN	SN	MH	TW
	KS	-				
	KN	0.003	-			
COI	SN	0.003	0.000	-		
	MH	0.015	0.019	0.019	-	
	TW	0.004	0.001	0.000	0.019	-
	KS	-				
	KN	0.008	-			
Cyt *b*	SN	0.008	0.001	-		
	MH	0.018	0.029	0.029	-	
	TW	0.008	0.001	0.000	0.029	-

Table 4. AMOVA of *T. jarbua* samples based on mtDNA sequences.

Source of variation	Sum of squares	Percentage of variation	F statistic	*P*
Among region (F_{CT})	800.958	39.52	0.3952	0.1896 ± 0.0134
Among populations within region (F_{SC})	11.1610	-0.14	-0.0033	0.0831 ± 0.0082
Within populations (F_{ST})	1476.92	62.13	0.3952	0.0000 ± 0.0000

Historical demography

The overall Tajima's *D* value was negative with an insignificant *p*-value, indicating deviation from evolutionary neutrality. Similarly, the Fu's Fs test which is based on the distribution of haplotypes, revealed negative but significant *p*-values for all five populations studied, indicating an excess of rare haplotypes or rare mutations in the population compared to what is expected under a neutral model of evolution. Following the results of Fu's Fs test, the hypothesis of neutral evolution was rejected.

In the present study, all populations demonstrated bimodal and ragged shaped patterns which points to the population having remained largely constant in size and that the lineage was widespread (Rogers and Harpending 1992). The scatterplot of the bimodal illustrations is shown in Figure 4. The results of mismatch distribution are contradictory to the results of the neutrality analysis. Hence, to further test the validity of the neutrality test results, we calculated the raggedness index and SSD under the demographic expansion model as shown in Table 5. *P*-

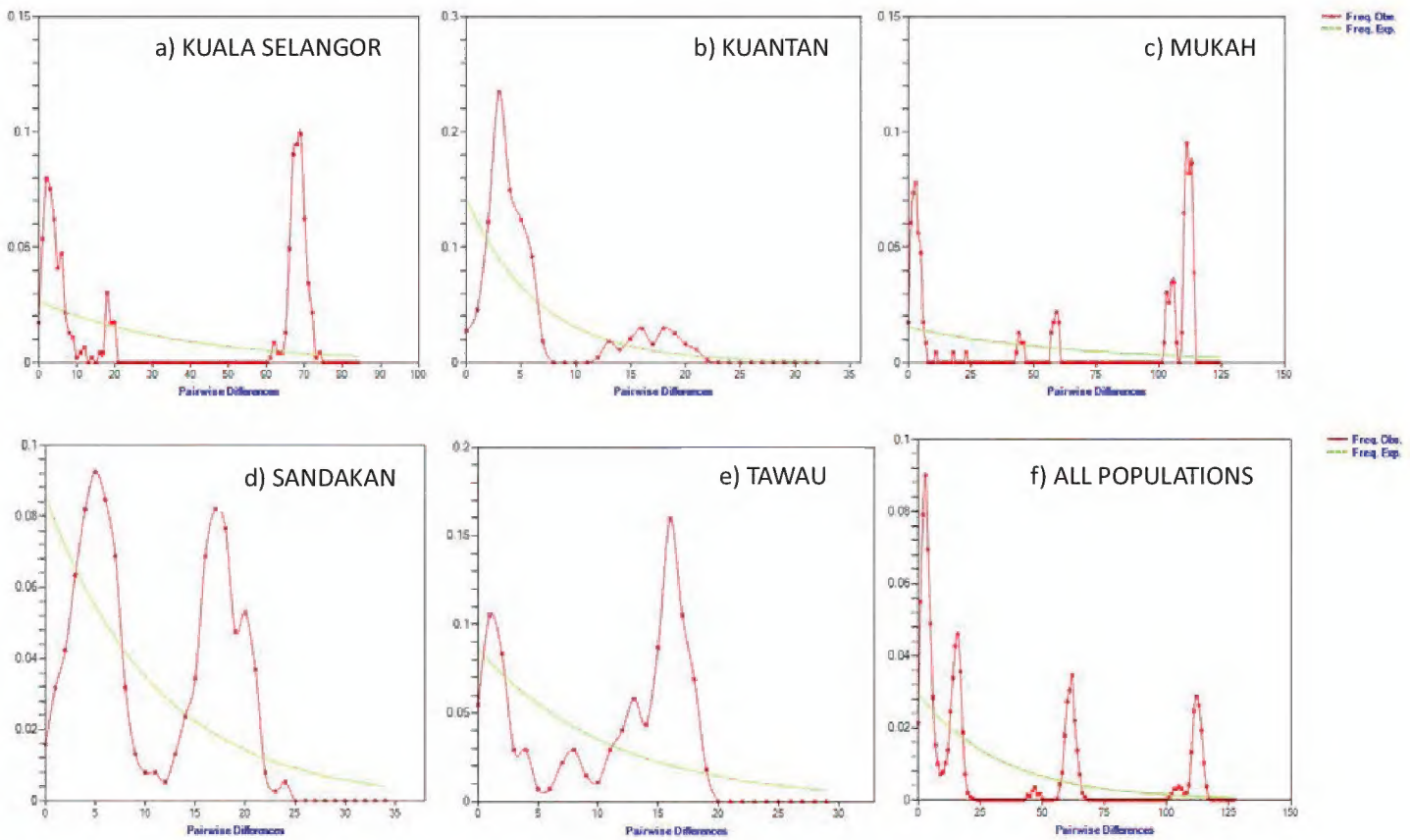

Figure 4. Pairwise number of difference (mismatch distribution) analysis was conducted using the constant population size model to observe the population size changes. The observed frequencies were represented by red dotted line. The frequency expected under the hypothesis of population expansion model was depicted by continuous green line. **a** Kuala Selangor **b** Kuantan **c** Mukah **d** Sandakan **e** Tawau **f** all populations.

Table 5. Parameter estimates of neutrality tests (Tajima's *D* statistic and Fu's Fs) and mismatch distribution (sum of squares deviation (SSD) and r = raggedness index) for each population. Significance (*p < 0.10) was determined using coalescent simulations.

	Neutrality test		Mismatch distribution		
	Tajima's *D*	**Fu's F_s**	**SSD**	**r**	**Curve**
KS	1.4817	-1.252	0.0296	0.0106	Bimodal
KN	-1.2595	-11.560	0.0113	0.0333	Bimodal
SN	-0.4265	-6.153	0.0301	0.0160	Bimodal
TW	1.0793	-2.075	0.0236	0.0266	Bimodal
MH	2.1863*	-1.093*	0.0288	0.0142	Bimodal
Total	-0.3132	-24.885*	0.0247	0.0201	Bimodal

values of SSD between the observed and expected mismatch distributions were all statistically insignificant ($p > 0.10$), indicating the presence of non-equilibrium and a population expansion event in *T. jarbua*. Besides that, the studied populations showed non-significant raggedness index ($p > 0.10$) indicating the data has relatively good fit to a model of population expansion (Harpending 1994). A ragged distribution suggests that the lineage was widespread (Excoffier et al. 1992; Rogers and Harpending 1992; Rogers 1995).

Discussion

Genetic diversity

Species identification was confirmed by morphological observation and DNA sequence data in which the intraspecific COI divergence was within the 2% threshold value (Hebert et al. 2003; Ward 2009; Srivathsan and Meier 2012). In general, a relatively high haplotype diversity (0.92–0.99) and low nucleotide diversity (0.03–0.34) were observed for the populations in this study. The combination of high haplotype diversity and low nucleotide diversity is common in pelagic marine fishes (Liu et al. 2015). This is likely due to rapid demographic expansion from a small effective population size, assuming there is sufficient time for the number of haplotypes to increase through mutation but insufficient for accumulation of large sequence differences (Grant and Bowen 1998; Avise 2000; Lowe et al. 2004). The inference of population expansion is further supported by the star-like patterns in Figure 3. Our results agree with an earlier study of *T. jarbua* populations from the Taiwanese waters which exhibited similarly high population haplotype diversity ranging from 0.86 to 1.00 as inferred from COI and Cyt *b* sequence data (Liu et al. 2015). In a study on *T. jarbua* populations from the Gulf of Aden, Lavergne et al. (2014) reported similarly high genetic diversities ranging from 0.216 to 0.698. The high number of haplotypes (83 haplotypes) in the present study is likely due to the high mutation rate of the mtDNA genes. Most of the haplotypes are unique to its region which may indicate the presence of different founding populations in the studied localities (Teixeira et al. 2011). The haplotype diversity of a relatively rapid evolving genome within a population often approaches 1.0 as many individuals will tend to have unique haplotypes (Freeland et al. 2011).

Genetic structure

A population's genetic structure is affected by genetic drift, local adaptation, and gene flow. In a marine environment, the development of population structure is greatly influenced by factors that affect dispersal, such as ocean currents, historical variance, and geographic distance coupled with differences in dispersal ability and habitat discontinuity (Saarman et al. 2010). Population structure inferred from mtDNA markers displays less genetic divergence in the pelagic and moderately pelagic species due to their potential to undertake long-distance migrations in oceanic waters (Jaafar 2014).

The haplotype tree (Fig. 2) revealed three major lineages but geographic structuring among the five populations is not distinct. In general, haplotypes specific to certain geographic regions did not form monophyletic groups, but appeared to be randomly distributed across the haplotype tree. Hap5 and Hap3 which recorded the highest distribution frequency are likely the ancestral haplotypes among the populations sampled.

Recent haplotypes were evolved directly or indirectly from the ancestral haplotypes. The existence of two ancestral points indicates that *T. jarbua* in Malaysia probably exist from two different sources. According to the coalescent theory, common haplotypes at the center of a network are inferred to be ancestral, while tip haplotypes at the periphery are derived or descendant from ancestral haplotypes (Akib et al. 2015). The occurrence of star-like patterns radiating from these major haplotypes suggests that *T. jarbua* populations have undergone significant population size expansions in the relatively recent past (Forster et al. 2001; Akib et al. 2015).

F_{ST} values are often used to infer gene flow, in which a lower F_{ST} value indicates low genetic divergence and higher gene flow. F_{ST} values below 0.05, as observed between SN and TW populations, indicate negligible genetic divergence, probably due to active exchange of genetic material between populations through breeding. Furthermore, the pairwise divergence between these populations is not statistically significant. According to Wright (1965), F_{ST} of 0–0.05 is described as little differentiation, 0.05–0.15 as moderate differentiation, 0.15–0.25 as great differentiation and values greater than 0.25 as very great differentiation. All populations studied showed moderate to very great pairwise differentiation except for TW-SN. The overall gene flow recorded was rather low (Nm = 0.82) which suggests limited genetic connectivity among the five populations.

Populations from the same region, i.e., TW and SN of Sabah, were the least genetically variable (Tables 2 and 3), which is likely due to the close geographical distance between these populations. The theory of fish migration across adjacent drainage systems due to flooding, which follows the one-dimensional stepping stone model that allows migration to adjacent population (Song et al. 2013), may apply in the case of *T. jarbua*. The significantly higher genetic differentiation between populations of KS-SN (pairwise F_{ST} = 0.2681) compared to KN-SN (pairwise F_{ST} = 0.0702) may be attributed to distance and physical barrier. Some genetic exchange can be expected since the Straits of Malacca connects the Andaman Sea and South China Sea via the narrow Tebrau strait. It is likely that mixing between the two bodies of water is very limited which supports the F_{ST} value obtained. However, the higher genetic variation between populations of MH-SN (pairwise F_{ST} = 0.5038) as compared to MH-KS (pairwise F_{ST} = 0.3310) implies that geographical distance is not the only driving factor of genetic variation among populations of *T. jarbua* in the Malaysian waters, similar to the results observed in the wider Gulf of Aden (Lavergne et al. 2014) where populations of adjacent locations showed low genetic connectivity despite the absence of a geographic barrier. Populations bordering a common origin such as the South China Sea (KN, MH, TW, and SN) may have evolved independently of each other over time, but there might have been insufficient time for genetic divergence to accumulate in these populations.

Another interesting finding of this study is the occurrence of shared haplotype between the populations from Peninsular and East Malaysia, India, Hainan, Philippines and Taiwan. Common haplotypes between localities and mixed haplotypes of different lineages in some populations in the current study can be explained by the biogeographical history of Southeast Asia (historically known as the Sundaland). Southeast

Asia is believed to have experienced simultaneous glaciation and consequent deglaciation along with its associated decrease and increase of seawater levels during the Pleistocene period, which greatly influenced continental and oceanic configuration (Voris 2000). The shared haplotypes between Malaysian populations and those from as far as India suggests that the range of population expansion after glacial retreat was not restricted to the South China Sea but also extended into the Indian Ocean (Liu et al. 2015). Lavergne et al. (2014) also reported high connectivity between populations in the Gulf of Aden and South China Sea due to the unique sharing of COI haplotypes between both regions. The haplotype sharing and their consequent gene flow may also be attributed to breeding migration, mutation, pelagic larvae, and sharing of common ancestors (Frankham 1996).

The MH population is the most genetically distinct with the highest between-group mean distances, haplotype and nucleotide diversity among the five populations. Geographical isolation of allopatric populations restricts gene flow between two populations, which in turn allows the evolution of a genome adapted to local condition (Hall 1993). Cluster I (MH) is estimated to form after the separation of Peninsular Malaysia from the Borneo Island due to the rise in the depth of the Sunda River between 40 to 100 m. This gradual separation was suspected to have caused accumulative genetic drift. According to Halliday (1993), genetic drift is likely to occur, particularly in small populations that are isolated from the main population and it may become the major source of genetic variation between some populations.

Among the four populations, MH is genetically closest to KS. Geological evidence suggests that the river systems of Sarawak were historically interconnected with most major river systems of Peninsular Malaysia via the Sunda River during Pleistocene glaciation (about 10000 years ago), thus allowing gene flow among these drainages (Kamarudin and Esa 2009). Gene flow from populations in the Straits of Malacca to those in Sarawak has been reported in several studies including Ryan and Esa (2006), Azhar and Hassan (2015), Samani et al. (2016) and Lau et al. (2018). Meanwhile, populations of SN-KN, which are separated by the South China Sea, showed high genetic connectivity (pairwise $F_{ST} = 0.0702$). This could probably be explained by the high migration ability of *T. jarbua* (>1000 km, Liu et al. 2015), human-mediated transfer through ballast waters (Liu et al. 2019) or past glaciation events. Furthermore, the high similarity in the sequence data (Table 3) perhaps indicates remnants of identical haplotypes from both populations, and that they were essentially similar at one time before the separation (Inger and Chin 2002). Eventually, sea level rise during the last Pleistocene period caused Borneo to be separated from mainland Asia (Peninsular Malaysia), which we suggest, resulted in shelf submergence and subsequent genetic differentiation between grunters from KN and SN. Pleistocene sea level fluctuations could also explain the incomplete divergence of grunters between East and Peninsular Malaysia. Similar evidence of a close genetic relationship between fishes of Borneo and mainland Asia in relation to their biogeographical history was discussed by several other authors (Pin et al. 2001; Nadiatul et al. 2011; Tan et al. 2012; Song et al. 2013).

Historical demography

Historical demographic expansions were determined by analysing the frequency distributions of pairwise differences between sequences (Rogers and Harpending 1992; Ray et al. 2003; Excoffier 2004). Neutrality tests with Tajima's *D* and Fu's Fs statistics estimate the deviation from neutrality, which is based on the expectation of a constant population size at mutation-drift equilibrium. Here, a negative Tajima's *D* signifies an excess of low frequency polymorphisms relative to expectation, indicating population size expansion or positive selection (Tajima 1983). The negative and significant Fu's Fs statistical value provides strong evidence for past population expansion, and rule out the possibility of genetic hitching or background selection, and evolutionary forces that produce a pattern similar to population expansion (Fu and Li 1993; Fu 1997; Okello et al. 2005). The *T. jarbua* populations displayed a genetic pattern typical of a population that has undergone a recent population expansion due to its two common haplotypes (Hap3 and Hap5) present across the range while the rest of the haplotypes are unique. The range expansion was a recent phenomenon and may not have achieved the migration-drift equilibrium, as shown by the lack of phylogeographical structure. Neutrality test statistics were in overall negatively significant and not consistent with a population at drift-mutation equilibrium.

The mismatch distribution is generally displayed as a multimodal pattern for populations showing demographic equilibrium. In contrast, a unimodal pattern depicts populations which have experienced recent demographic expansion (Rogers and Harpending 1992). In the results, all localities presented a multimodal pattern proving recent expansion. The hypothesis that the observed data fit the sudden expansion model was tested using the SSD and the raggedness index. Here, non-significant values for SSD signifies that the observed data do not deviate from that expected under the model of expansion. Non-significant raggedness index also indicates population expansion. Our observations of non-significant values in goodness-of-fit distribution for all populations suggest that population expansion occurred recently (Rogers and Harpending 1992).

Conclusion

To summarize, we found 1) high haplotype diversity but low nucleotide diversity among *T. jarbua* populations in Malaysia; 2) significant results suggesting population expansion of *T. jarbua* in this region; 3) despite the three genetic clusters observed in the haplotype tree and median-joining network, no obvious population structuring was detected among geographically distinct populations. Common haplotypes among populations and haplotypes from several populations in each genetic cluster indicate high genetic connectivity among the populations. This study assesses the genetic diversity and population structure of *T. jarbua* in Malaysia for appropriate conservation and management strategies. Conservation of crescent grunter at its natural variation level is required as it forms a diverse group of taxa with 83 haplotypes distributed across Malaysia. The haplotype composition surveyed in the present study may provide a baseline for future comparisons to monitor the temporal variability of haplotype frequency and population structure. This study also has indirectly revealed the

dispersal power of *T. jarbua* through its high mobility and rapid adaptability to a newly colonized area. Further studies can be conducted using larger sample size and temporal replicates, samples collected from other areas of geographical distributions, and sequence data from other mtDNA genes or information based on nuclear DNA. This research contributed useful data for future large scale biogeographical and taxonomic studies of this species.

Animal ethics

The fish species that was employed in this study is not categorized as endangered species under the IUCN list and all the samples were collected from fish markets and landing sites.

Acknowledgements

This study was supported by the University of Malaya, Research University Grant (RU009E-2018), Top 100 Universities in The World Fund (TU001-2018), IF030B-2017; Ministry of Science and Technology (108-2119-M-110-005) and the China-ASEAN Maritime Cooperation Fund project "Monitoring and conservation of the coastal ecosystem in the South China Sea". We would also like to thank Surajwaran Mangaleswaran, an English professional for checking on the language used in this paper.

References

Akib NAM, Tam BM, Phumee P, Abidin MZ, Tamadoni S, Mather PB, Nor SAM (2015) High connectivity in *Rastrelliger kanagurta*: influence of historical signatures and migratory behaviour inferred from mtDNA cytochrome b. PLoS One 10(3): e0119749. https://doi.org/10.1371/journal.pone.0119749

Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, 447pp.

Azhar MAAM, Hassan R (2015) Population genetics of Tor Douronensis in Sarawak–a Revisit. Borneo Journal of Resource Science and Technology 5(2): 1–15. https://doi.org/10.33736/bjrst.218.2015

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution 16: 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Borsa P, Durand JD, Chen WJ, Hubert N, Muths D, Mou-Tham G, Kulbicki M (2016) Comparative phylogeography of the Western Indian Ocean reef fauna. Acta Oecologica 72: 72–86. http://doi.org/10.1016/j.actao.2015.10.009

Canoy RJC, Quilang JP (2015) Molecular Phylogeny of Philippine Tigerperches (Perciformes: Terapontidae) Based on Mitochondrial Genes. [Unpublished]

Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. Nature Reviews Genetics 10: 195. https://doi.org/10.1038/nrg2526

Dahanukar N, Kaymaram F, Alnazry H, Al-Husaini M, Almukhtar M, Hartmann S, Alam S, Sparks JS (2017) *Terapon jarbua*. The IUCN red list of threatened species 2019. https://www.iucnredlist.org/species/166892/46643542

Department of Fisheries Malaysia (2009) Valid local name of Malaysian marine fishes. Department of Fisheries Malaysia. Ministry of Agriculture and Agro-based Industry, 180 pp.

Dowling DK, Friberg U, Lindell J (2008) Evolutionary implications of non-neutral mitochondrial genetic variation. Trends in Ecology & Evolution 23: 546–554. https://doi.org/10.1016/j.tree.2008.05.011

Du J, Loh K-H, Hu W, Zheng X, Amri AY, Ooi JLS, Ma Z, Rizman-Idid M, Chan AA (2019) An updated checklist of the marine fish fauna of Redang Islands, Malaysia. Biodiversity Data Journal 7: e47537. https://doi.org/10.3897/BDJ.7.e47537

Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. Molecular Ecology 13: 853–864. https://doi.org/10.1046/j.1365-294X.2003.02004.x

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491.

Frankham R (1996) Relationship of genetic variation to population size in wildlife. Conservation Biology 10(6):1500–1508.

Freeland JR, Kirk H, Peterson SD (2011) Molecular Ecology, 2nd ed. Wiley-Blackwell, 449 pp.

Froese R, Pauly D (2018) FishBase. http://fishbase.sinica.edu.tw/Summary/SpeciesSummary.php?id=4458 [August 2019]

Forster P, Torroni A, Renfrew C, Röhl A (2001) Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. Molecular Biology and Evolution 18: 1864–1881. https://doi.org/10.1093/oxfordjournals.molbev.a003728

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147(2): 915–925.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genetics 133(3): 693–709.

Golani D, Appelbaum-Golani B (2010) First record of the Indo-Pacific fish the *Jarbua terapon* (*Terapon jarbua*) (Osteichthyes: Terapontidae) in the Mediterranean with remarks on the wide geographical distribution of this species. Scientia Marina 74(4): 717–720.

Grant W, Bowen BW (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. Journal of Heredity 89: 415–426. https://doi.org/10.1093/jhered/89.5.415

Hall M (1993) Species, speciation and extinction evolution – A Biological and paleontological approach. Addison-Wesley Publishing Company, Great Britain, 391pp.

Hall T, Biosciences I, Carlsbad C (2011) BioEdit: an important software for molecular biology. GERF Bull Biosci 2: 60–61.

Halliday T (1993) Natural selection In: Skelton P (Ed.) Evolution - A biological and palaeontological approach. Addison-Wesley Publishing Company, Great Britain, 141–142.

Harpending H (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Human Biology 66(4): 591–600. https://doi.org/10.2307/41465371

Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. Current Anthropology 34: 483–496. https://doi.org/10.1086/204195

Hebert PD, Ratnasingham S, De Waard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London Series B: Biological Sciences 270: S96–S99.

Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. Genetics 132(2): 583–589.

Hyde J, Lynn E, Humphreys Jr R, Musyl M, West A, Vetter R (2005) Shipboard identification of fish eggs and larvae by multiplex PCR, and description of fertilized eggs of blue marlin, shortbill spearfish, and wahoo. Marine Ecology Progress Series 286: 269–277. https://doi.org/10.3354/meps286269

Inger RF, Chin PK (2002) Fresh-water fishes of North Borneo. Natural History Publications, Borneo, 346 pp.

Jaafar TNAM (2014) DNA barcoding and population genetic structure of Malaysian marine fishes. PhD Thesis, Bangor University, Bangor. 327 pp.

Kamarudin KR, Esa Y (2009) Phylogeny and phylogeography of *Barbonymus schwanenfeldii* (Cyprinidae) from Malaysia inferred using partial cytochrome b mtDNA gene. Journal of Tropical Biology & Conservation 5: 1–13.

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33: 1870–1874. https://doi.org/10.1093/molbev/msw054

Lau JS, Ransangan J, Rodrigues KF (2018) Genetic diversity and population structure of the Asian Green Mussel (Pernaviridis) in the waters of Sabah, Malaysia based on mitochondrial DNA D-Loop Sequences. Turkish Journal of Fisheries and Aquatic Sciences 18(1): 109–117. https://doi.org/10.4194/1303-2712-v18_1_12

Lavergne E, Calvès I, Meistertzheim AL, Charrier G, Zajonz U, Laroche J (2012) Estuarine fish biodiversity of Socotra Island (NW Indian Ocean): from the community to the functioning of *Terapon jarbua* populations. PhD Thesis, University of Western Brittany.

Liu L, Zhang X, Li C, Zhang H, Yanagimoto T, Song N, Gao T (2019) Population genetic structure of Marbled Rockfish, *Sebastiscus marmoratus* (Cuvier, 1829), in the northwestern Pacific Ocean. ZooKeys 830: 127. https://doi.org/10.3897/zookeys.830.30586

Liu SYV, Huang IH, Liu MY, Lin HD, Wang FY, Liao TY (2015) Genetic stock structure of *Terapon jarbua* in Taiwanese waters. Marine and Coastal Fisheries 7: 464–473. https://doi.org/10.1080/19425120.2015.1074966

Lowe A, Harris S, Ashton P (2004) Genetic diversity and differentiation – Ecological genetics: design, analysis, and application. Blackwell Publishing Oxford, 50–105. https://doi.org/10.1093/aob/mci073

Miu TC, Lee SC, Tzeng WN (1990) Reproductive biology of *Terapon jarbua* from the estuary of Tamshui River. Journal of the Fisheries Society of Taiwan 17(1): 9–20. https://doi.org/10.29822/JFST.200812.0004

Musarrat UAin RYF, Masood Z (2015) Gonadosomatic index of a Teraponid Species, *Terapon jarbua* (Forsskal, 1775) (Family: Teraponidae) of Karachi Coast, Pakistan. International Journal of Biology and Biotechnology 12(4): 575–578.

Nadiatul H, Daud S, Siraj S, Sungan S, Moghaddam F (2011) Genetic diversity of Malaysian indigenous Mahseer, *Tor douronensis* in Sarawak river basins as revealed by cytochrome c oxidase I gene sequences. Iranian Journal of Animal Biosystematics 7(2): 119–127.

Nandikeswari R, Sambasivam M, Anandan V (2014) Estimation of fecundity and gonadosomatic index of *Terapon jarbua* from Pondicherry Coast, India. International Journal of Biological, Biomolecular, Agricultural, Food and Biotechnological Engineering 8(1): 61–65. https://doi.org/10.5281/zenodo.1090709

Nei M (1987). Molecular Evolutionary Genetics. Columbia University Press, 512 pp.

Okello JBA, Nyakaana S, Masembe C, Siegismund HR, Arctander P (2005) Mitochondrial DNA variation of the common hippopotamus: evidence for a recent population expansion. Heredity 95(3): 206–215. https://doi.org/10.1038/sj.hdy.6800711

Okumuş İ, Çiftci Y (2003) Fish population genetics and molecular markers: II-molecular markers and their applications in fisheries and aquaculture. Turkish Journal of Fisheries and Aquatic Sciences 3: 51–79.

Pin LC, Teen LP, Ahmad A, Usup G (2001) Genetic diversity of *Ostreopsis ovata* (Dinophyceae) from Malaysia. Marine Biotechnology 3: 246–255. http://doi.org/10.1007/s101260000073

Rao DV, Devi K, Rajan P (2000) Account of Ichthyofauna of Andaman & Nicobar Islands, Bay of Bengal. Zoological Survey of India 48p.

Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. Molecular Biology and Evolution 20: 76–86.

Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. Evolution 49: 608–615. https://doi.org/10.1111/j.1558-5646.1995.tb0229.x

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Molecular Biology and Evolution 9: 552–569. https://doi.org/10.1093/oxfordjournals.molbev.a040727

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496–2497. https://doi.org/10.1093/bioinformatics/btg359

Ryan JR, Esa YB (2006) Phylogenetic analysis of Hampala Fishes (Subfamily Cyprininae) in Malaysia inferred from partial mitochondrial cytochrome B DNA Sequences. Zoological Science 23(10): 893–902. https://doi.org/10.2108/zsj.23.893

Saarman NP, Louie KD, Hamilton H (2010) Genetic differentiation across eastern Pacific oceanographic barriers in the threatened seahorse *Hippocampus ingens*. Conservation Genetics 11: 1989–2000. http://doi.org/10.1007/s10592-010-0092-x

Samani NK, Esa Y, Amin SN, Ikhsan NFM (2016) Phylogenetics and population genetics of *Plotosus canius* (Siluriformes: Plotosidae) from Malaysian coastal waters. PeerJ 4: e1930. http://doi.org/10.7717/peerj.1930

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics 152: 1079–1089.

Shyama SDC, Lim PE, Poong SW, Du J, Loh KH (2020) Relationships between Sagittal otolith size and fish size of *Terapon jarbua* (Teleostei, Terapontidae) in Malaysian waters. Journal of Oceanology and Limnology http://doi.org/10.1007/s00343-019-9193-7

Slatkin M (1987) Gene flow and the geographic structure of natural populations. Science 236: 787–792. https://doi.org/10.1126/science.3576198

Song LM, Munian K, Abd Rashid Z, Bhassu S (2013) Characterisation of Asian snakehead murrel *Channa striata* (Channidae) in Malaysia: an insight into molecular data and morphological approach. The Scientific World Journal 2013: e917506. https://doi.org/10.1155/2013/917506

Srivathsan A, Meier R (2012) On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. Cladistics 28: 190–194.

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. Genetics 105: 437–460.

Tajima F (1989) The effect of change in population size on DNA polymorphism. Genetics 123: 597–601.

Tan MP, Jamsari AFJ, Azizah MNS (2012) Phylogeographic pattern of the striped snakehead, *Channa striata* in Sundaland: ancient river connectivity, geographical and anthropogenic singnatures. PLoS One 7: e52089. https://doi.org/10.1371/journal.pone.0052089

Teixeira S, Cambon-Bonavita MA, Serrão EA, Desbruyeres D, Arnaud-Haond S (2011) Recent population expansion and connectivity in the hydrothermal shrimp *Rimicaris exoculata* along the Mid-Atlantic Ridge. Journal of Biogeography 38: 564–574. https://doi.org/10.1111/j.1365-2699.2010.02408.x

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22: 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. Journal of Biogeography 27: 1153–1167. https://doi.org/10.1046/j.1365-2699.2000.00489.x

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD (2005) DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B: Biological Sciences 360: 1847–1857. https://doi.org/10.1098/rstb.2005.1716

Ward RD (2009) DNA barcode divergence among species and genera of birds and fishes. Molecular Ecology Resources 9(4): 1077–1085. https://doi.org/10.1111/j.1755-0998.2009.02541.x

Whitehead A, Anderson SL, Kuivila KM, Roach JL, May B (2003) Genetic variation among interconnected populations of *Catostomus occidentalis*: implications for distinguishing impacts of contaminants from biogeographical structuring. Molecular Ecology 12: 2817–2833. https://doi.org/10.1046/j.1365-294X.2003.01933.x

Wright S (1931) Evolution in Mendelian populations. Genetics 16: 1–97.

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19: 395–420.

Wu G, Wu C, Wang Q, Luo J (2016) The complete mitochondrial genome of the *Terapon jarbua* (Perciformes: Terapontidae). Mitochondrial DNA Part A 27: 3430–3431. https://dx.doi.org/10.3109/19401736.2015.1022734

Zaya DN, Molano-Flores B, Feist MA, Koontz JA, Coons J (2017) Assessing genetic diversity for the USA endemic carnivorous plant *Pinguicula ionantha* RK Godfrey (Lentibulariaceae). Conservation Genetics 18: 171–180. https://dx.doi.org/10.1007/s10592-016-0891-9

Appendix I

Percentage of nucleotide composition based on populations.

ID	Nucleotide composition (%)																	
	COI						Cyt b						Combine					
	A	T	C	G	A + T	C + G	A	T	C	G	A + T	C + G	A	T	C	G	A+T	C+G
KS	22.4	28.8	30.9	17.9	51.2	48.8	23.5	28.7	32.9	14.9	52.2	47.8	22.9	28.9	32.0	16.2	51.8	48.2
KN	22.6	28.9	30.7	17.8	51.5	48.5	23.5	28.7	32.8	14.9	52.2	47.7	23.0	29.0	31.9	16.1	52.0	48.0
MH	22.5	28.4	31.4	17.8	50.9	49.2	23.7	28.4	32.9	15.0	52.1	47.9	23.0	28.6	32.2	16.2	51.6	48.4
SN	22.6	28.9	30.7	17.8	51.5	48.5	23.5	28.7	32.8	14.9	52.2	47.7	23.0	29.0	31.9	16.1	52.0	48.0
TW	22.6	28.9	30.8	17.8	51.5	48.6	23.5	28.7	32.9	14.9	52.2	47.8	23.0	28.9	32.0	16.1	51.9	48.1
Total	22.2	29.2	30.9	17.7	51.4	48.6	23.6	28.7	32.9	14.9	52.3	47.8	23.0	28.9	32.0	16.1	51.9	48.1
HAI	22.5	28.9	30.7	17.8	51.4	48.5	23.4	28.8	32.9	14.8	52.2	47.7	23.1	28.9	32.0	16.1	52.0	48.1
IND	22.3	29.5	30.5	17.7	51.8	48.2	23.6	28.8	32.8	14.8	52.4	47.6	22.9	29.7	31.6	15.8	52.6	47.4
PHI	22.9	28.9	30.7	17.5	51.8	48.2	23.4	28.6	32.9	15.1	52.0	48.0	23.2	28.7	32.0	16.1	51.9	48.1
TAI	22.5	28.9	30.7	17.8	51.4	48.5	23.6	28.7	32.9	14.8	52.3	47.7	23.1	28.8	32.0	16.1	51.9	48.1

Appendix 2

Polymorphic site analysis based on COI, Cyt *b* and combined gene. C, conserved site; V, variable site; Pi, parsimony informative sites; S, singleton sites.

ID	COI (631 bp)				Cyt *b* (815 bp)				Combine (1446 bp)			
	C	V	Pi	S	C	V	Pi	S	C	V	Pi	S
KS	602	29	20	9	742	73	53	20	1344	102	73	29
KN	618	13	6	7	793	22	14	8	1411	35	20	15
MH	575	56	54	2	716	99	95	4	1291	155	149	6
SN	617	14	7	7	787	28	20	8	1404	42	27	15
TW	619	12	9	3	798	17	14	3	1417	29	23	6
Total	557	74	60	14	687	128	111	17	1244	202	171	31

Appendix 3

Sequence data used in this study. All data were downloaded from NCBI.

Location	Accession number		Publication
	COI	Cyt *b*	
Hainan	NC027281	NC027281	Wu et al. 2016
Taiwan	KP204162	KP152133	Liu et al. 2015
Kochi, India	KC774674	KC774717	Lenka et al. 2014 (Unpublished)
Philippines	KF999840	KF999856	Canoy and Quilang 2015 (Unpublished)
Malaysia	MN529663–MN529796	MN529797–MN529930	This study

Appendix 4

Frequency distribution of haplotypes according to localities. Highlighted columns indicate shared haplotypes.

Haplotype	Total	KS	KN	TW	SN	MH	PHI	TAI	HAI	IND
Hap_1	3	3								
Hap_2	1	1								
Hap_3	17	4	4	7		1				1
Hap_4	1	1								
Hap_5	21	5	6	1	5	3		1		
Hap_6	1	1								
Hap_7	1	1								
Hap_8	2	1	1							
Hap_9	1	1								
Hap_10	1	1								
Hap_11	1	1								
Hap_12	1	1								
Hap_13	1	1								
Hap_14	1	1								
Hap_15	1	1								
Hap_16	1	1								
Hap_17	1	1								
Hap_18	1	1								
Hap_19	1	1								
Hap_20	1	1								
Hap_21	1	1								
Hap_22	1	1								
Hap_23	1		1							
Hap_24	4		4							

Haplotype	Total	KS	KN	TW	SN	MH	PHI	TAI	HAI	IND
Hap_25	1		1							
Hap_26	1		1							
Hap_27	1		1							
Hap_28	1		1							
Hap_29	1		1							
Hap_30	1		1							
Hap_31	1		1							
Hap_32	1		1							
Hap_33	1		1							
Hap_34	1		1							
Hap_35	1		1							
Hap_36	1		1							
Hap_37	1		1							
Hap_38	1		1							
Hap_39	1				1					
Hap_40	1				1					
Hap_41	5				5					
Hap_42	1				1					
Hap_43	1				1					
Hap_44	1				1					
Hap_45	2			1	1					
Hap_46	1				1					
Hap_47	1				1					
Hap_48	1				1					
Hap_49	1				1					
Hap_50	1				1					
Hap_51	1				1					
Hap_52	1				1					
Hap_53	1				1					
Hap_54	1				1					
Hap_55	1				1					
Hap_56	1				1					
Hap_57	1				1					
Hap_58	2					2				
Hap_59	1					1				
Hap_60	1					1				
Hap_61	1					1				
Hap_62	1					1				
Hap_63	1					1				
Hap_64	2					2				
Hap_65	1					1				
Hap_66	1					1				
Hap_67	1					1				
Hap_68	1					1				
Hap_69	2					2				
Hap_70	1					1				
Hap_71	1					1				
Hap_72	1			1						
Hap_73	2			2						
Hap_74	1			1						
Hap_75	1			1						
Hap_76	3			3						
Hap_77	1			1						
Hap_78	1			1						
Hap_79	1			1						
Hap_80	1			1						
Hap_81	1			1						
Hap_82	1			1						
Hap_83	1			1						
Hap_84	1								1	
Hap_85	1						1			